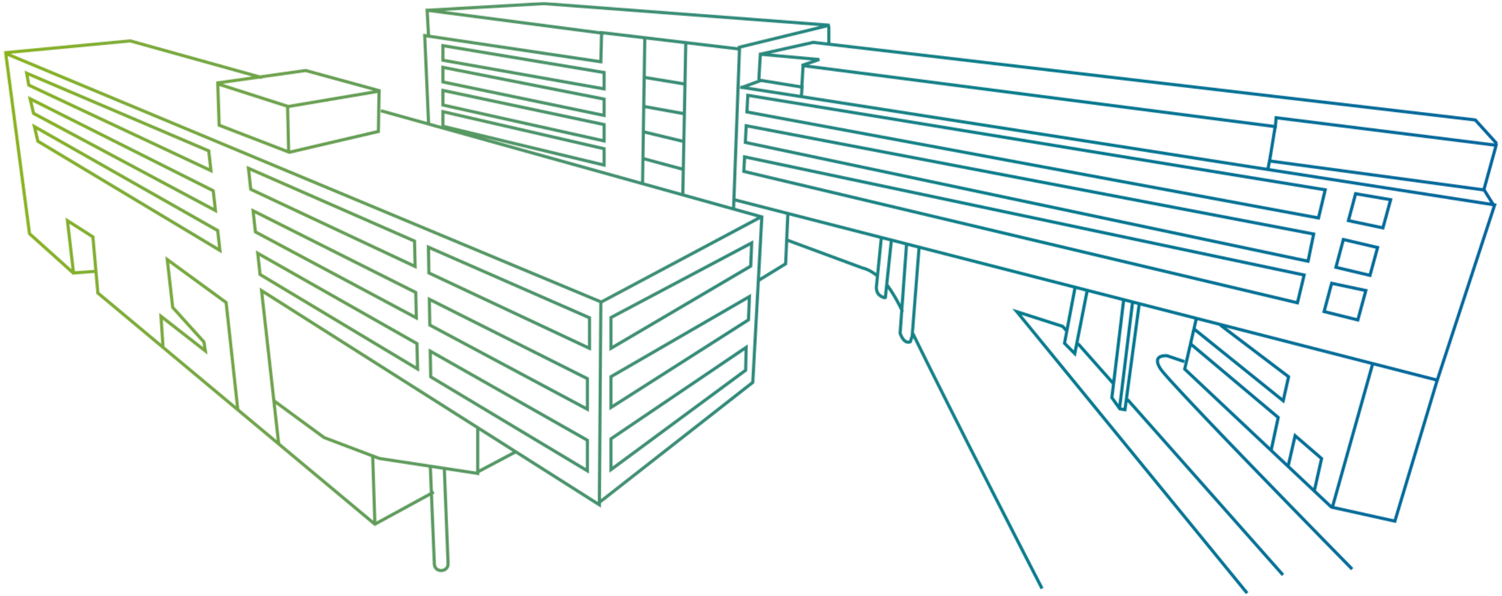

VAIDS-DSML exercise block 9

Jasmin Duvivié

simple linear regression

- problem

- 1599 samples of wines
- 11 features
- 1 target: quality
- 1 feature as linear regressor: alcohol

- solution

- use sklearn.linear_model
- construct model manually
- compare ols parameters (slope and intercept)

simple linear regression

function for linear regression

$$y_i = \beta_0 + \beta_i x_i + \varepsilon_1$$

as matrices

$$\boldsymbol{y} = \boldsymbol{x\beta} + \boldsymbol{\varepsilon}$$

dimensions of those matrices in simple linear regression

$$(n \times 1) = (n \times 2)(2 \times 1) + (n \times 1)$$

coefficients

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$$

estimate for $\boldsymbol{\beta}$

$$\boldsymbol{b} = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$$

prediction $\hat{y}$

$$\hat{\boldsymbol{y}} = \boldsymbol{xb}$$

simple linear regression

error $\boldsymbol{e}$

$$\boldsymbol{e} = \boldsymbol{y} - \widehat{\boldsymbol{y}}$$

or

$$\boldsymbol{e} = \boldsymbol{y} - \boldsymbol{x}\boldsymbol{b}$$

note dimensions

$$(n \times 1) = (n \times 1) - (n \times 1)$$

remember the goal is to minimize the error:

$$\min \sum_{i}^{n} e_i^2$$

or

$$\min \boldsymbol{e}^T\, \boldsymbol{e}$$

or

$$\min(\boldsymbol{y} - \boldsymbol{x}\boldsymbol{b})^T\, (\boldsymbol{y} - \boldsymbol{x}\boldsymbol{b})$$

note dimensions

$$(1 \times n)(n \times 1) = (1 \times 1)$$

simple linear regression

we want to find the b which minimizes the error, so we take the partial derivative with regard to b and set it to 0

$$\frac{\partial}{\partial \mathrm{b}}(\boldsymbol{y}-\boldsymbol{x}\boldsymbol{b})^T(\boldsymbol{y}-\boldsymbol{x}\boldsymbol{b})=\mathbf{0}$$

draw the transposition into the parentheses and remember to reverse the terms if you transpose a product

$$\frac{\partial}{\partial \mathrm{b}}(\boldsymbol{y}^T-\boldsymbol{b}^T\boldsymbol{x}^T)(\boldsymbol{y}-\boldsymbol{x}\boldsymbol{b})=\mathbf{0}$$

aus-x-en

$$\frac{\partial}{\partial \mathrm{b}}\boldsymbol{y}^T\boldsymbol{y}-\boldsymbol{y}^T\boldsymbol{x}\boldsymbol{b}-\boldsymbol{b}^T\boldsymbol{x}^T\boldsymbol{y}+\boldsymbol{b}^T\boldsymbol{x}^T\boldsymbol{x}\boldsymbol{b}=\mathbf{0}$$

simple linear regression manually

$$\frac{\partial}{\partial \mathrm{b}} \boldsymbol{y}^T \boldsymbol{y} - \boldsymbol{y}^T \boldsymbol{x} \boldsymbol{b} - \boldsymbol{b}^T \boldsymbol{x}^T \boldsymbol{y} + \boldsymbol{b}^T \boldsymbol{x}^T \boldsymbol{x} \boldsymbol{b} = 0$$

note the dimensions of only this part of the derivative

$$-\boldsymbol{y}^T \boldsymbol{x} \boldsymbol{b}$$

$$(1 \times n)(n \times 2)(2 \times 1)$$

$$(1 \times n)(n \times 1)$$

$$(1 \times 1)$$

so it is a scalar

and we have that two times in our term

$$-\boldsymbol{y}^T \boldsymbol{x} \boldsymbol{b} - \boldsymbol{b}^T \boldsymbol{x}^T \boldsymbol{y}$$

$$-2\boldsymbol{b}^T \boldsymbol{x}^T \boldsymbol{y}$$

so the whole term looks like this now

$$\frac{\partial}{\partial \mathrm{b}} \boldsymbol{y}^T \boldsymbol{y} - 2\boldsymbol{b}^T \boldsymbol{x}^T \boldsymbol{y} + \boldsymbol{b}^T \boldsymbol{x}^T \boldsymbol{x} \boldsymbol{b} = \boldsymbol{0}$$

simple linear regression

now derive: terms with no b disappear, terms with a b loose the b, terms with b^2 become $2b$

$$-2\boldsymbol{x}^T\boldsymbol{y} + 2(\boldsymbol{x}^T\boldsymbol{x})\boldsymbol{b} = \boldsymbol{0}$$

solve for b (the osl parameter estimates)

$$2(\boldsymbol{x}^T\boldsymbol{x})\boldsymbol{b} = 2\boldsymbol{x}^T\boldsymbol{y}$$

$$(\boldsymbol{x}^T\boldsymbol{x})\boldsymbol{b} = \boldsymbol{x}^T\boldsymbol{y}$$

$$(\boldsymbol{x}^T\boldsymbol{x})^{-1}(\boldsymbol{x}^T\boldsymbol{x})\boldsymbol{b} = (\boldsymbol{x}^T\boldsymbol{x})^{-1}(\boldsymbol{x}^T\boldsymbol{y})$$

$$(\boldsymbol{x}^T\boldsymbol{x})^{-1}(\boldsymbol{x}^T\boldsymbol{x}) = 1$$

$$\boldsymbol{b} = (\boldsymbol{x}^T\boldsymbol{x})^{-1}(\boldsymbol{x}^T\boldsymbol{y})$$

check the dimensions one last time

$$(2 \times 1) = (2 \times n)(n \times 2)(2 \times n)(n \times 1)$$

$$(2 \times 1) = (2 \times 2)(2 \times 1)$$

same as

$$\hat{\boldsymbol{y}} = \boldsymbol{x}\boldsymbol{b}$$

yay, checks out :)